LLNL-AR-402633

The Earth System Grid Center for Enabling Technologies (ESG-CET):

Scaling the Earth System Grid to Petascale Data

Climate simulation data are now securely accessed, monitored, cataloged, transported, and distributed to the national and international climate community

Semi-Annual Progress Report for the Period
October 1, 2007 through March 31, 2008

Table of Contents

1 Executive Summary

This report, which summarizes work carried out by the ESG-CET during the period October 1, 2007 through March 31, 2008, includes discussion of overall progress, period goals, highlights, collaborations and presentations. To learn more about our project, please visit the Earth System Grid website. In addition, this report will be forwarded to the DOE SciDAC project management, the Office of Biological and Environmental Research (OBER) project management, national and international stakeholders (e.g., the Community Climate System Model (CCSM), the Intergovernmental Panel on Climate Change (IPCC) 5th Assessment Report (AR5), the Climate Science Computational End Station (CCES), etc.), and collaborators.

The ESG-CET executive committee consists of David Bernholdt, ORNL; Ian Foster, ANL; Don Middleton, NCAR; and Dean Williams, LLNL. The ESG-CET team is a collective of researchers and scientists with diverse domain knowledge, whose home institutions include seven laboratories and one university: Argonne National Laboratory (ANL), Los Alamos National Laboratory (LANL), Lawrence Berkeley National Laboratory (LBNL), Lawrence Livermore National Laboratory (LLNL), National Center for Atmospheric Research (NCAR), Oak Ridge National Laboratory (ORNL), Pacific Marine Environmental Laboratory (PMEL), and University of Southern California, Information Sciences Institute (USC/ISI). All work is accomplished in close collaboration with the project's stakeholders and domain researchers and scientists.

1.1 Overall goal for this reporting period

During this semi-annual reporting period, the ESG-CET increased its efforts on completing software investigations, component prototyping, and software development—transitioning from requirements gathering, architecture, and design to implementation, in order to meet our crucial testbed deadline. As we strove to complete software for our testbed, we continued to provide production-level services to the community. These services continued for IPCC AR4, CCES, and CCSM, Parallel Clilmate Model (PCM), Parallel Ocean Program (POP), and Cloud Feedback Model Intercomparison Project (CFMIP), and they were extended to include the North America Regional Climate Change Assessment Program (NARCCAP) data.

1.2 Highlights

1.2.1 LLNL ESG Portal Highlights

The Program for Climate Model Diagnosis and Intercomparison (PCMDI) at LLNL depends on the DOE SciDAC sponsored Earth System Grid (ESG) to publish and disseminate the Coupled Model Intercomparison Project (CMIP3) Intergovernmental Panel on Climate Change (IPCC) 4th Assessment Report (AR4) data to the community. Extracting science and knowledge from this large data archive, this ESG database supports a specific user community of over 1,800. These data provided the basis for over 400 publications and played an important role in support of the 2007 report by the IPCC, recently named a co-recipient of the 2007 Nobel Peace Prize for its work in "disseminating greater knowledge about man-made climate change". Figure 1 shows the daily download rate over time.

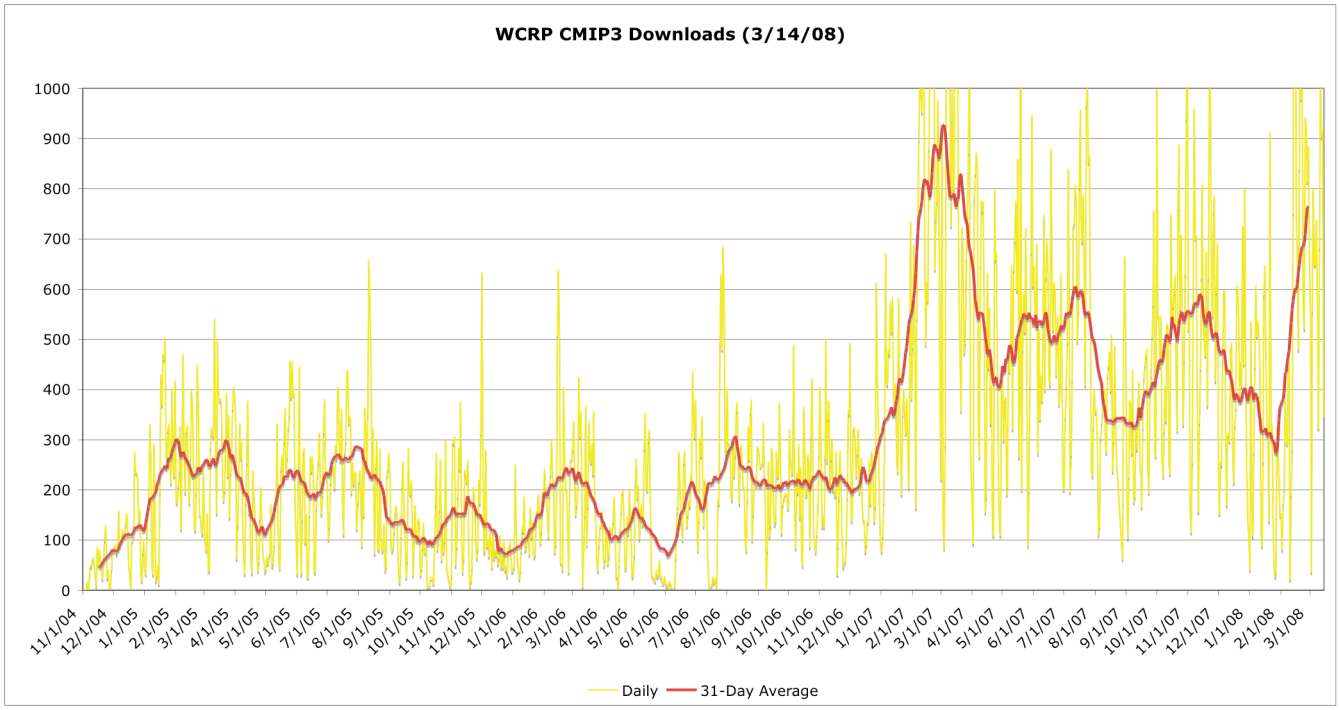

Figure 1: CMIP3 (IPCC AR4) Download Rates in Gigabytes/day

1.2.2 ORNL ESG Portal Highlights

Ongoing operation and maintenance of ESG portal for the Carbon Land Model Intercomparison Project (C-LAMP) continues.

ORNL implemented a new data node ("ESGtg") to replace “sleepy”. ESGtg is located in ORNL's TeraGrid network enclave, giving it direct 10 Gbps network links to NCAR and ANL, with possible links to other ESG institutions in the future. The new node also has significantly more local disk storage, which will facilitate large-scale data movement. This work involved significant contributions from the LBNL team to port their Storage Resource Manager (SRM) software to the new 64-bit hardware platform, as well as from the ISI team for debugging Replica Location Storage (RLS) issues.

1.2.3 LANL ESG Node Highlights

LANL converted and transferred 250 GB of new high-resolution ocean simulations to their data node “oceans11”. The NCAR data portal is now working to publish these simulations. LANL is also helping to design the next generation of ESG so that publication occurs—as much as possible—on the data node side. For ESG-CET Data Producers, minimal interaction with the ESG-CET Gateway is desirable when publishing new datasets.

LANL also upgraded the “Tomcat” server and related software on “oceans11”. Some recent tests from “oceans11” had download failures (unrelated to the software upgrade), which are in the process of being resolved.

1.2.4 LBNL Storage Resource Manager Highlights

LBNL received a special request from a ORNL climate modeling user (Marcia Branstetter) to download 2,000 files (each about 60 MB in size) from ESG. The current setup was limited to 100 files, and a total of 10 GB of data. Both the Portal and DataMover-Lite (DML) had to be upgraded to deal with hundreds

of files simultaneously. In collaboration with the Scientific Data Management (SDM) center, a student from North Carolina State university (Paul Breimyer) set up a java program that communicated directly with the portal and DML software, avoiding the web interfaces. This required the upgrade of a command-line version of DML, a task performed by Alex Sim, of LBNL. This tool has been successfully implemented, and the user is grateful for the timely help.

1.2.5 PMEL Product Delivery Services Highlights

For the ESG project in year one, the Live Access Server (LAS) was converted into a generalized workflow engine and was distributed to other ESG partners. Collaborative work based upon this foundation will continue for the duration of the project. Version 7 of the product server implements the LAS service request protocol (XML) for delivering "information products" to end-users and to other tiers of the ESG-CET system—typically visualizations, tables, and file subsets. It can call upon a number of important "back end services" and link them into useful workflows. These services include relational databases (SQL via JDBC), netCDF file IO, OPeNDAP-CF (for use with curvilinear multidimensional grids, including aggregation services), the PMEL-developed Ferret application, the PCMDI-developed CDAT application for graphics rendering; and the OPeNDAP-DAPPER protocol for access to collections of time series and profile observations. Also added was the ability to output data via the OPeNDAP data access protocol (in addition to the previously available input), and of gridded output using the OGC Web Coverage Service (WCS). PMEL also has implemented a powerful server-side computation capability that can perform functions essential to the numerical model output datasets that are the focus of ESG-CET—regridding, evaluation of mathematical expressions, basic statistics (e.g. averaging, finding extrema, variances, etc.); and data filters (smoothers, gap-fillers, etc.)

In recent work, LAS has added output capabilities in the OGC Web Mapping Service (WMS—for latitude/longitude maps); and input capabilities for *in-situ* observational data, using application schemas within the OGC Web Feature Service (WFS) specification. A large proportion of our recent development efforts have been devoted to server-side support for highly interactive browser-based user interfaces—an API of AJAX-callable services that provide dataset metadata (variable names, dimensionality of variables, coordinate ranges; menu items for user selection of coordinates, etc.). We have participated in project-wide work to define the distributed architecture of ESG product services and have begun work incorporating candidate access control frameworks into LAS.

1.2.6 ANL Security, Data, and Services Highlights

ANL has continued to work on designing federated security architecture for the ESG-CET project, with recent focus on solutions for single sign-on and automatic trust root provisioning. We have gathered security requirements and use cases from the ESG-CET team and external collaborators to design a federated solution for user authentication and single sign-on capabilities.

In the last few months, ANL has evaluated various solutions for providing a single sign-on solution for web servers within the ESG-CET infrastructure. We have conducted a detailed evaluation of the OpenID protocol and implementation to ascertain viability for use in ESG-CET project. We are in the process of evaluating Shibboleth, another single sign-on solution as a comparable alternative.

In addition to solutions for web servers, ANL is also leading the effort on designing single sign-on solutions for applications that do not use HTTP or HTTPS protocol. We evaluated use of the PKI X509 infrastructure with Online CA solution for leveraging other application servers such as GridFTP and OpenDAP, and we demonstrated the solution using a test installation of MyProxy Online CA. Work is in progress to integrate this solution with the ESG-CET Gateway software.

ANL has also worked on solutions for scalable automated provisioning of trust roots in the ESG-CET infrastructure. The proposed solution, using the MyProxy infrastructure, will allow for administrators to maintain trust roots in centralized repository and ensure that the resources in ESG-CET and clients are automatically updated with the most recent information.

ANL continues to work on evaluating server-side processing solutions using LAS, and on the publishing of data. This work is in collaboration with Environment Science Division at ANL, with a focus on publishing their datasets for remote sharing, and providing server-side processing of the data.

2 Overall Progress

During this reporting period, progress was made in the key areas that are necessary to meet ESG-CET objectives, goals and milestones.

2.1 ESG-CET Gateway

The first prototype of the ESG-CET Gateway ("alpha" version) was completed in January 2008 and was deployed on a testbed at NCAR immediately thereafter (see Figure 2). The Gateway represents the central component of the next generation ESG-CET 3-tier architecture (which will also include Global Services and Data Nodes), and it will serve the following purposes:

- Hosting critical services such as data search and browsing, user registration, user management, metrics, authentication and authorization, metadata persistence, user workspace, etc.
- Acting as broker for data requests sent to Data Node components (data download, transfer from deep storage, subsetting, and visualization).
- Exposing all user interface components.

The architecture of the Gateway alpha prototype is illustrated in Figure 3. We expect that initially three Gateways (and possibly more) will be deployed as part of the first ESG-CET system (at NCAR, ORNL and LLNL/PCMDI), and will be connected with a large number of Data Nodes to physically serve the data to the users.

Specifically, the following functionality was included in the Gateway alpha prototype:

- **User and Group management:** generic ESG-CET registration, application for specific group membership (for example: ÇMIP5 (IPCC AR5), CCSM, CES, NARCCAP), group administration, root-level operations;
- **User authentication:** implemented either via local Gateway authentication provider, or via separate Central Authentication Server (CAS);
- **Data browsing:** hierarchical browsing of datasets and data files are driven from the relational domain model database;
- **URL-based Access Control:** any page in the system, or any portion of a page, can be restricted to those users having specific group permissions;
- **Data Download:** includes browser-driven download of full data files from a remote Data Node, or from a local disk (with the Data Node application server and the Gateway sharing a web Single Sign On session though the CAS server). Also includes scriptable (wget) download of full data files from a remote Data Node, or from a local disk (with security enforced via limited lifetime authorization tokens, both for Gateway files and files served from the Data Node);

- **Semantic Data Search:** entails faceted search based on RDF triple store repository, which is populated from metadata harvested from the domain model relational database. The Gateway alpha includes two prototype search interfaces for evaluation:
    - A "modal" interface that uses a fixed large number of facets. As the user selects one of the facets, the options available in all others change accordingly.
    - A "hierarchical browsing" interface which allows the user to select which facets to use in the search, and in which order. The user selection for each facet determines the range of available options for the next facet. In the future, the user could save his/her preference for search categories and find them when next accessing this page.
- **Data Visualization:** a Gateway-side visualization user interface that embeds LAS UI components, and executes requests to an LAS Product Server hosted on a remote Data Node.

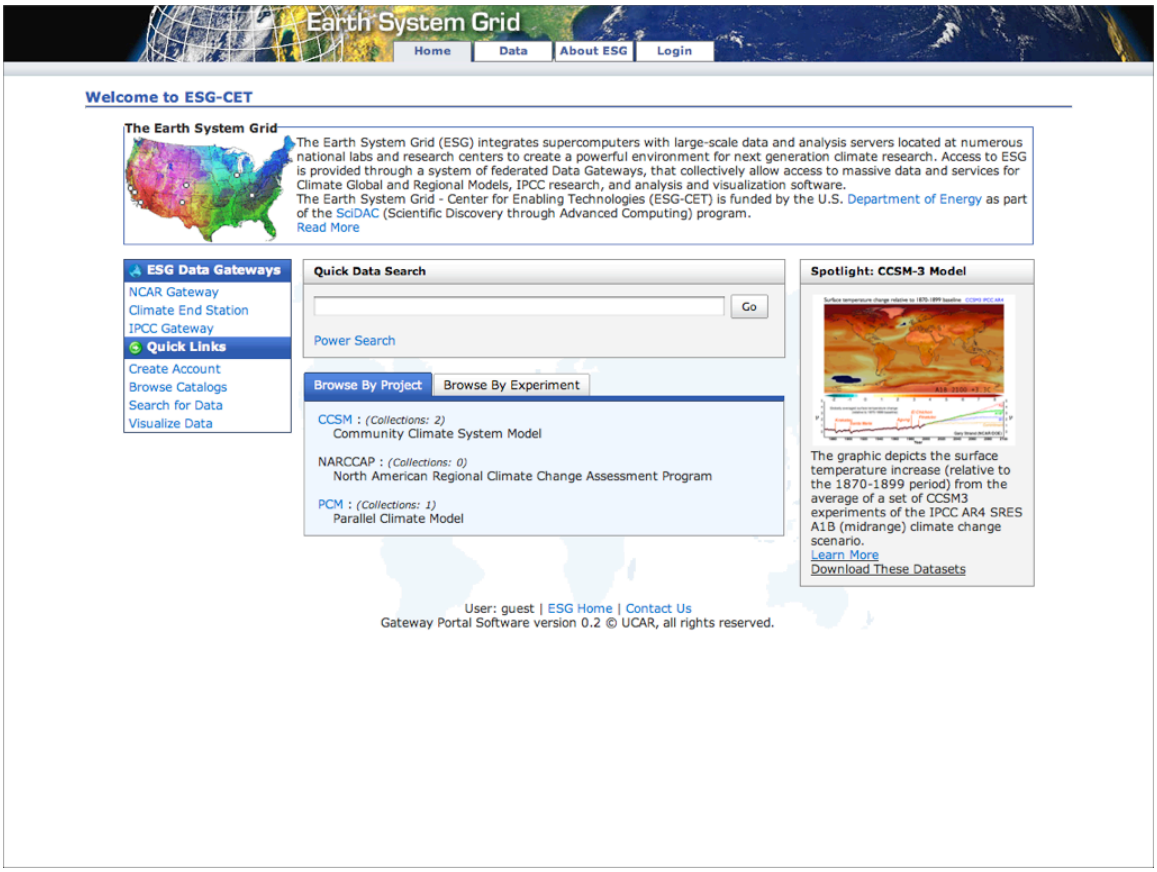

Figure 2: The home page of the Gateway alpha web portal.

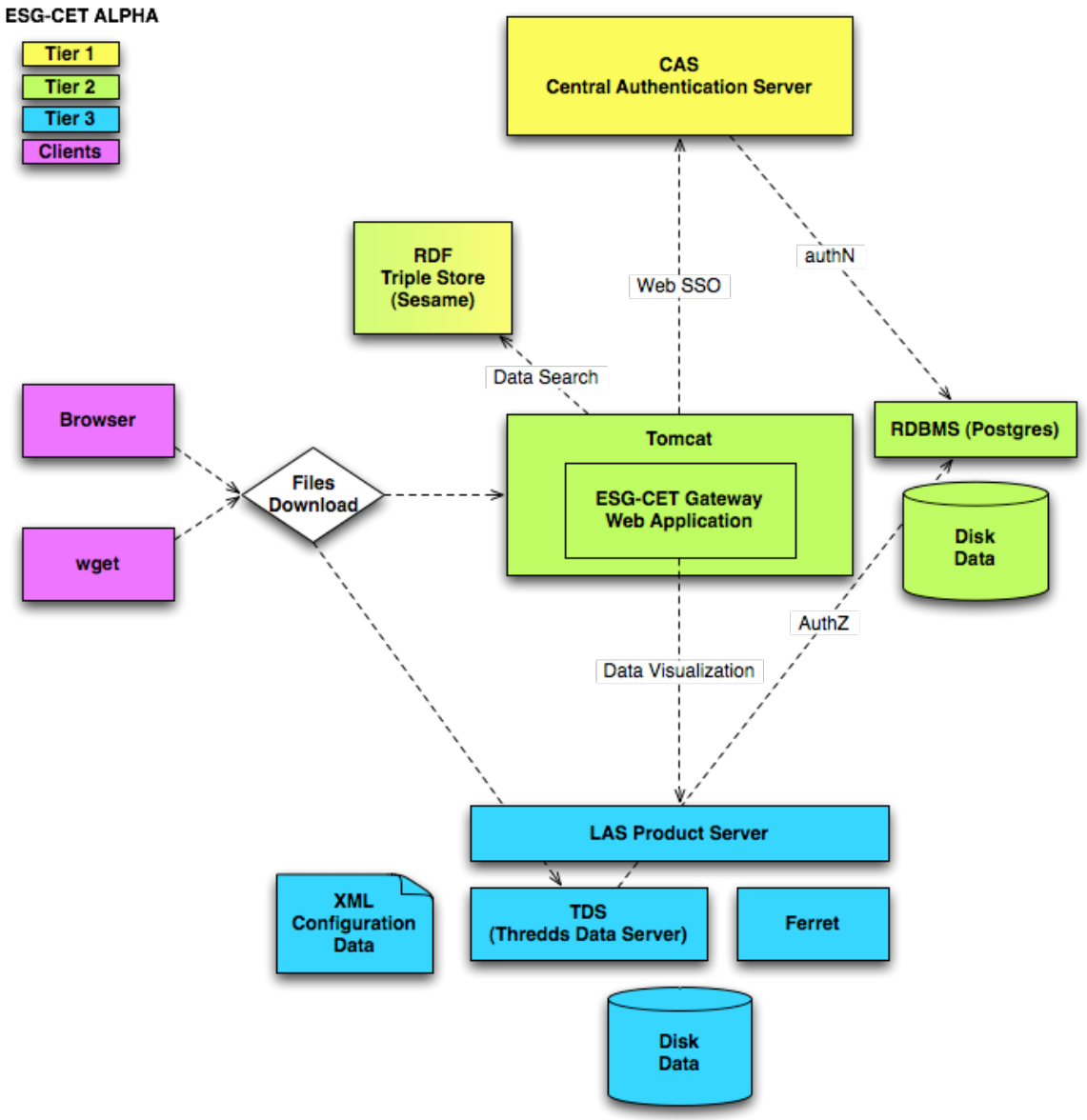

Figure 3: The 3-tier architecture of the Gateway alpha prototype.

2.2 ESG-CET Data Node: Publisher

Like the ESG-CET Gateway, the Data Node development is also underway with the design and implementation of the Data Publisher. Integration of the Data Publisher and the Gateway is scheduled to occur in the third quarter of 2008. The goal of the ESG-CET Data Node team is to provide powerful, clean, and simple node configuration and publishing tools that are lightweight and easily installed.

Data producers will have the ability to execute a standalone ESG publication tool like the one shown in Figure 4. The Data Publisher operation state will either be: “on-line” or “off-line”. When “on-line” the Data Publisher will attempt to connect to tape or rotating disk storage severs. While “off-line” no such connection to a storage device is attempted. For each dataset, the specification may be a list of files, a directory, or a Unix style regular expression.

Extraction information is retrieved from the dataset metadata and is transferred to the node metadata where quality-control checking of data accuracy occurs. If there are errors, then the original data can be rescanned for correction immediately, or the publisher can refrain from correcting the error until a later time. In the latter case, the level of associated metadata drops to a very basic level that specifies only the filename and size.

Once dataset extraction and quality control are performed, the data producer signals the Gateway to harvest the published data. This action will allow users to find and access newly published data through the Gateway's faceted search selection capabilities for data downloads or visualization. See Figure 5 for a visual depiction of these steps.

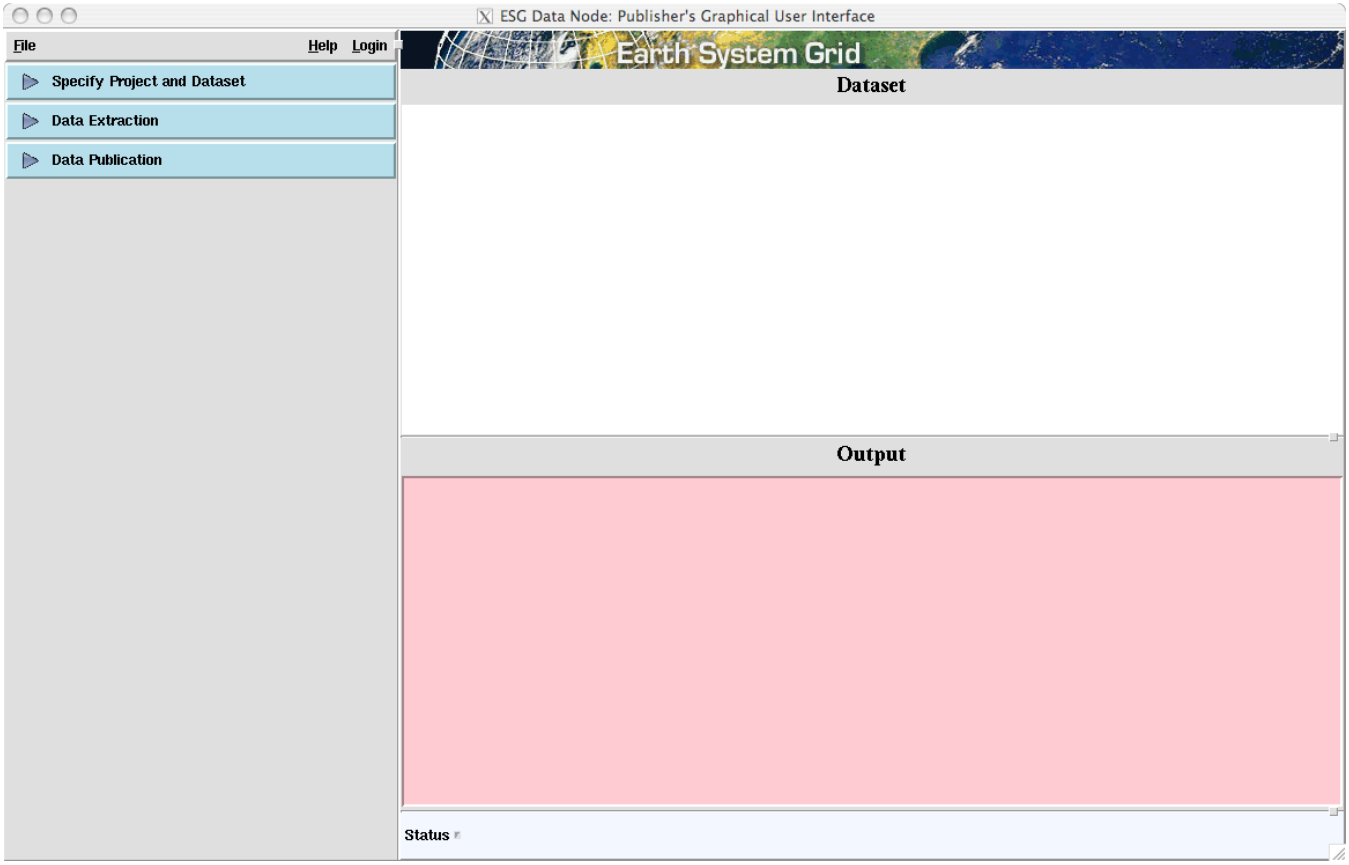

Figure 4: The ESG Data Node "Publisher's" client prototype.

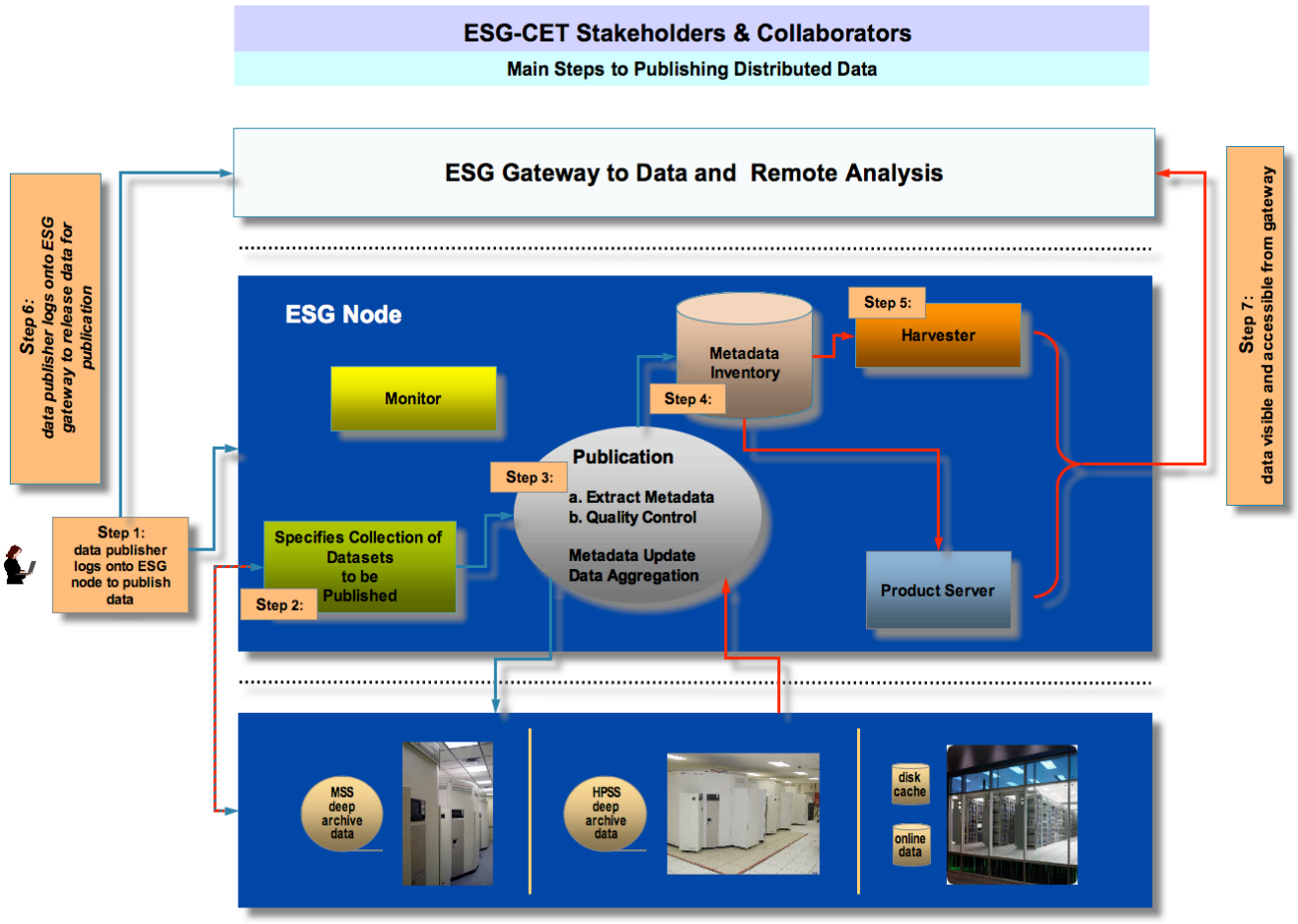

Figure 5: Steps to publishing data from the ESG Data Node

2.3 ESG-CET Cyber Security

Secure access to data and resources plays a crucial role in the ESG-CET. The security model must safeguard data, resources, and the credentials of both users and services—but without placing an undue burden on the users. The main challenge is to strike the right balance between ensuring the required security level of the overall system and promoting its practical utility.

Additionally, the scope of ESG continues to enlarge with the requirement to federate additional national and foreign sites (e.g. the Princeton Geophysical Fluid Dynamics Laboratory (GFDL), the British Atmospheric Data Center (BADC), and the University of Tokyo Center for Climate System Research, Japan). The associated use cases translate into a requirement for a single sign-on solution for the browser clients as well as the web service and GridFTP clients.

The overall ESG security architecture must be flexible enough to accommodate site-specific needs of individual groups, as well as general infrastructure needs. Toward this end, we have focused on the evaluation and prototyping of single sign-on solutions for web browsers and services, and for GridFTP clients.

See the following URLs for additional details:

- http://www.ci.uchicago.edu/wiki/bin/view/FranksProjects/OpenIDEvalESG
- http://www-unix.mcs.anl.gov/~ranantha/esg/PKISSO.html

2.4 Product Services

The ESG-CET is intended to serve customers of a broad spectrum of sophistication. These users will range from numerical modelers who want access to "raw" model output files and verbatim subsets of model output, to climate impacts investigators who want rapid access to these data without the complexities of model-specific coordinate systems, to those users who only want to quickly visualize the overall behaviors of models. The petascale nature of the ESG data holdings require that significant levels of data reduction take place at the server in order to satisfy these customers – both through straightforward subsetting and decimation and through specific analysis operations, such as the computing of spatio-temporal averages. In the ESG-CET architecture, we refer to the steps that convert raw data into analysis results and visualizations as "product services".

As described in the *"PMEL Product Delivery Services Highlights"* section, the Live Access Server (LAS) has been extended into a generalized workflow engine for the creation and delivery of ESG products. A service-oriented approach in which "back-end services" are accessed via SOAP has been employed in order to make the architecture adaptable to the range of products that it must provide. New output product capabilities were added including dynamic (adaptive, zoomable) Google Earth® visualizations; on-the-fly custom animations; and a "slide sorter" tool that allows scientists to inspect and to difference gridded fields.

Recent work has allowed users to interact directly with LAS output by using mouse events to zoom, or to request further products. AJAX services also have been used to create a highly interactive user interface that allows scientists to transform-on-the-fly, and to visualize data. Collaboration is proceeding with NCAR developers to embed these capabilities into the ESG portal.

2.5 Storage Resource Managers (SRM)

One of the important requirements over the last few months was the adaptation of the SRM running at ORNL to a new 64-bit machine. This version of the SRM (called HRM- Hierarchical Storage Manager because it accesses HPSS) had to be modified, and extensive scalability tests were successfully conducted.

In anticipation of the new ESG-CET gateways and data nodes, the new version of SRM, called BeStMan (Berkeley Storage Manager) will be deployed. BeStMan is a new more flexible implementation that facilitates interfacing with a variety of storage systems. Work over the last reporting period included the enhancements made to BeStMan to interact with the versions of HPSS at ORNL and NERSC, as well as the mass storage system at NCAR, call MSS). Plan for incorporating BeStMan with the newly designed Gateway are underway.

2.6 DataMover-Lite

A new DML UI was developed – designed to support a concurrent file transfer algorithm for more efficient file transfer performance. This is accomplished by using multiple threads (usually about 10), each responsible for initiation and monitoring of a single file. Thus, even if the underlying file transfer mechanism is slow (for example, using LAHFS - Lightweight Authorized HTTP File Server), the

concurrent multiple transfers can take advantage of the network bandwidth, and increase the aggregate transfer bandwidth.

Another enhancement to DML was the development of a new version that runs on the IBM SP–a requirement for users at the University of Illinois at Urbana-Champaign (UIUC).

2.7 Monitoring

The ISI team continues to provide the monitoring services infrastructure that allows ESG to detect and repair component failures. These monitoring services are essential for the reliable operation of the ESG portals and services. When failure messages occur, the ISI team helps to identify the causes of failures in the ESG infrastructure.

In this reporting period, support for the ESG monitoring infrastructure included several efforts. We helped the ORNL team set up the monitoring probes for the SRM, TRM, and RLS services running on the new ESGtg machine. We also assisted the ORNL and LBNL teams in tracking down problems with SRM servers that were failing on several sites—mainly at ORNL, but also at NCAR and LBNL. Finally, the monitoring team is also doing ongoing work to integrating dataportal's monitoring framework with the Globus Toolkit Version 4.1 codebase. In particular, the goal of this work is to capture information on transitions of services from up to down and the reverse.

2.8 Support for Current Replica Location Server (RLS) Deployments in ESG

The ISI team has provided ongoing support for RLS catalog servers that are deployed in the current ESG infrastructure. During this quarter, that support has involved tracking and debugging required for RLS servers at LANL, LLNL and NCAR. In addition, ISI staff helped ORNL to troubleshoot a problem with the new RLS server deployed on the ESGtg machine, and to synchronize the new RLS catalog with the old one on machine "sleepy".

2.9 Role of RLS in New ESG Architecture

The ISI team led a detailed discussion and email thread about the future role of the RLS in the next-generation ESG architecture. As a result of this discussion, ISI planned to investigate the utility of RLS catalogs for replicating location mappings among gateways. As part of this work, Rob Schuler of ISI worked with NCAR to extract mappings from their gateway database and prepare these for insertion in an RLS catalog. In the reporting period ahead, ISI plans to demonstrate the utility of the RLS by replicating these mappings on gateway nodes at ORNL, NCAR and LLNL.

2.10 Extending RLS State Sharing Mechanisms for Sharing RDF Triples Among Gateways

One outcome of the discussion of the future role of RLS was the identification of the RLS state sharing mechanism, which uses soft state update techniques, as an approach that might be generally useful for sharing state among ESG gateways. In particular, the ESG team is seeking a service that can extract Resource Description Framework (RDF) triples from a store on one gateway, and upload these triples to a store at another gateway. The RLS team has begun to investigate the design of such a service. Important design questions include the necessary size of these updates, and whether they can be compressed. In the coming reporting period, ISI plans to design this service and provide an initial prototype.

2.11 Investigation of Accounting and Auditing Technologies on Other Grid Projects

A final effort by the ISI team this reporting period was to survey the mechanisms used for auditing and accounting in three major grid projects (OSG, TeraGrid and CEDPS). The results of this survey were reported to ESG-CET, with the hope that some of these techniques may prove useful supplements to the logging and accounting already performed. We also identified accounting and logging issues that ESG-CET has not yet addressed, such as how accounting data gathered at nodes will be communicated to gateways.

2.12 Deployment of ESG-CET

In order to address software engineering, deployment, and maintenance issues for the operational ESG-CET, ORNL has begun to evaluate the Virtual Data Toolkit (VDT), an Open Science Grid product, as the basis on which the ESG-CET software stack might be built, tested, and deployed on operational ESG-CET nodes. To this end, the ORNL ESGtg server was built using VDT software packages wherever possible. We also held extensive discussions with the OSG team about adding software to VDT and/or creating adjunct packages for ESG-CET software.

3 Architectural Design Diagrams, Requirement Documents and Use Cases

All architectural design diagrams and requirement and use case documents referenced in Section 2 of this report can be viewed on the ESG-CET website.

4 ESG-CET Group Meetings

The ESG-CET executive committee holds weekly conference calls each Tuesday at 10:00 a.m. PDT. These meetings discuss priorities and issues that make up the agenda for the weekly project meetings held via the AccessGrid (AG) every Thursday at 12:00 p.m. PDT. At these meetings, the entire team discusses project goals, design and development issues, technologies, timelines, and milestones. Given the need for more in-depth conversation and examination of work requirements, the following upcoming face-to-face meetings will be held during the next reporting period:

4.1 ESG-CET Executive Meeting

In early-April, the ESG-CET executive committee will convene at the Office of Advanced Scientific Computing Research (OASCR) Computer Science Principal Investigator Meeting, to be held in Denver, CO to discuss among other things the upcoming ESG-CET All-hands meeting. The meetings will also covered project management, technical direction, collaborations, and overall project direction.

4.2 Architecture and Development ESG-CET All-hands Meeting

In late-April, the ESG-CET team will come together over a three-day period for the ESG Prototyping All-hands Meeting. The meeting will be held at the National Center for Climate Research (NCAR) in Boulder, CO. We will discuss interactions with science stockholders and the subprojects cross collaborations, as well as planning a path forward to meet the end-of-year ESG-CET testbed deadline.

5 Collaborations

To effectively build an infrastructure capable of dealing with petascale data management and analysis, we established connections with other funded DOE Office of Science SciDAC projects and programs at

various meetings and workshops, such as the SciDAC 2007 Conference held in Boston, MA. In particular, collaborations were established with the following groups:

5.1 Hybrid Coordinate Ocean Model (HyCOM) consortium (NOAA, Navy, et. al.)

NOAA/PMEL (Steve Hankin, ESG co-PI) is a partner in the Hybrid Coordinate Ocean Model (HyCOM) consortium [http://hycom.rsmas.miami.edu/]. The HyCOM Consortium has developed a high- resolution (1/12 degree) operational, global ocean modeling capability under cooperative US Navy and NOAA funding. The HyCOM model presents unique technical challenges, owing to its complicated coordinate system and large data volumes, but the needs of HYCOM overlap in many respects with those of the ocean components of the IPCC AR5 climate models. There is thus a significant and productive two-way transfer of technical capabilities developed in support of ESG and HyCOM.

The ESG-CET collaboration has worked towards enabling support, within the current ESG operational system, for publishing and distributing NARCCAP (North America Climate Regional Climate Change Project) data. An extensive data management plan was developed that involves distributed data access from the ESG portal at NCAR to data resources stored at both NCAR and PCMDI. The existing user registration system was extended to allow a separate community of NARCCAP users vetted by specific administrators, and the first test users were approved for access.

5.2 NOAA Geophysical Fluid Dynamics Laboratory

The NOAA GFDL Fluid Dynamics Laboratory is an active contributor to AR5 and an active participant in the ESG SciDac. V. Balaji [Head, GFDL Modeling Systems Group] is a frequent participant and active contributor in ESG teleconferences and meetings, resulting in a vigorous bi-directional exchange of ideas and technology. NOAA/PMEL (Steve Hankin, ESG co-PI) shares an MOU with GFDL for the development of the Laboratory's data portal, thereby also implementing an active two-way technology transfer between NOAA and ESG.

5.3 Global Organization for Earth System Science Portal (GO-ESSP)

The GO-ESSP is a collaboration designed to develop a new generation of software infrastructure that will provide distributed access to observed and simulated data from the climate and weather communities. Of the seven members of the GO-ESSP steering committee, three are members of the ESG-CET team: Steve Hankin, Don Middleton, and Dean N. Williams.

5.4 Earth System Curator (ESC)

The ESG-CET and the Earth System Curator (ESC) are working together to develop prototype ontology, user interface, and relational databases to include additional information on the model configurations that produce datasets. ESG-CET team members Luca Cinquini and Don Middleton (as a co-PI) are working closely with this group. Other ESG-CET team members may be involved as work progresses.

5.5 Scientific Data Management (SDM) Center for Enabling Technology (SciDAC CS CET)

Based on the experience with DataMover-Lite, a new client version of an SRM (known as "SRM-Lite") was developed in order to invoke it to move files to and from sites that have one-time-password (OTP) security or other highly secure systems. SRM-Lite is similar in design to DML, but has only a

command-line interface. We plan to use this tool in a workflow system that the SDM center uses, called Kepler, as well as other application projects.

5.6 VACET: VisTrails

VisTrails is a new scientific workflow management system. The LLNL team is currently working to modify the Climate Data Analysis Tools (CDAT) XML output to allow the use of VisTrails. If successful, this workflow may be implemented in the next generation of ESG architecture.

5.7 Institute for Ultrascale Visualization

Jian Huang (U. Tennessee), a member of the SciDAC Institute for Ultrascale Visualization, is developing tools for web-based collaborative visualization of climate data, with initial application to data from the ORNL C-LAMP portal. ORNL plans to deploy Huang's visualization server on the ORNL C-LAMP computer, and will work towards integrating it as a "product server" in the next-generation ESG architecture. Huang has been invited to present a paper on this work at the May 2008 International Symposium on Collaborative Technologies and Systems (CTS 2008).

6 Outreach, Papers, Presentations and Posters

List of talks and posters presented during this time period:

6.1 ASCAC-BERAC Charge Committee

Ian Foster and Dean Williams served on the Joint Advanced Scientific Computing Advisory Committee (ASCAC) and the Biological and Environmental Research Advisory Committee (BERAC) Subcommittee to investigate barriers and bottlenecks to achieving successful outcomes of complementary investments—specifically in climate modeling by ASCR and BER. Ian and Dean contributed to a 19-page report that was transmitted to Ray Orbach, Director of the DOE Office of Science.

6.2 Paper in the Cyberinfrastructure Technology Watch (CTWatch)

D. N. Williams, D. E. Bernholdt, I. T. Foster, and D. E. Middleton, 2007: *The Earth System Grid Center for Enabling Technologies: Enabling community access to petascale climate datasets*. *Cyberinfrastructure Technology Watch (CTWatch) Quarterly*, Vol. 3, number 4.

6.3 Paper in the Bulletin of the American Meteorological Society (BAMS)

The following article by ESG-CET participants is a follow-on to the Meehl et al. 2007 BAMS article on *the CMIP3 multi-model dataset*s that were an important basis for the IPCC AR4 effort:

D. N. Williams, R. Ananthakrishnan, D. E. Bernholdt, S. Bharathi, D. Brown, M. Chen, A. L. Chervenak, L. Cinquini, R. Drach, I. T. Foster, P. Fox, D. Fraser, J. Garcia, S. Hankin, P. Jones, C. Kesselman, D. E. Middleton, J. Schwidder, R. Schweitzer, R. Schuler, A. Shoshani, F. Siebenlist, A. Sim, W. G. Strand, and N. Wilhelmi, 2008: The Earth System Grid: Enabling access to multi-model climate simulation data. *Bulletin of the American Meteorological Society* (in review).

6.4 International Symposium on Collaborative Technologies and Systems

W. Kendall, M. Glatter, J. Huang, F. Hoffman, and D. E. Bernholdt, 2008: Web enabled collaborative climate visualization in the Earth System Grid. Invited paper for the 2008 International Symposium on Collaborative Technologies and Systems (CTS 2008), to appear.

6.5 LLNL 2007 Computation Directorate Annual Report

A.A. Mirin, and D.N. Williams, 2008: Climate modeling – Conducting research and providing enabling technology. *LLNL 2007 Computation Directorate Annual Report*.

6.6 Global Organization for Earth System Science Portal (GO-ESSP) Mini-Workshop

In mid-October, Steve Hankin, Don Middleton, and Dean N. Williams attended the GO-ESSP mini-workshop which focused on establishing open access to distributed climate data archives.

6.7 SDM Center All-hands Meeting

Hosted by Arie Shoshani, Don Middleton and Dean Williams attended the SDM Center All-hands Meeting where climate analysis and workflow use cases where discussed. Don and Dean represented ESG-CET and made technology collaboration contacts (e.g., for Kepler and VisTrails).